



Вариант №2 ETL пайплайны в NoSQL

Цели практики

Освоить способ построения ETL пайплайна в **Mongo**.

Описание задания:

Способ построения ETL пайплайна в Mongo.

- Для того, чтобы приступить к выполнению задания, необходимо, используя образ Mongo, аналогично тому как показано на уроке установить в Docker образ сервера БД.
- Для работы с Mongo рекомендуем использовать <https://robomongo.org/>, у которого есть расширенная версия на 30 дней или Free.
- Есть массив данных пользователей в CSV формате https://disk.yandex.ru/d/YagnK1O_qaicOw.

Создать коллекцию с документами в Mongo с указанными полями:

- username
- country
- comes_from
- hobbies
- ip_address
- birthday.

Составить запрос, который:

а) Первый - выводит пользователей старше 01/01/2010. Вывод должен содержать все поля, кроме hobbies.

б) Второй - должен выполнять то же самое, но в агрегационном пайплайне.

По итогу выполнения задания у вас должны быть текстовые файлы с тремя работоспособными запросами и два ответа от БД по двум запросам последнего пункта этого задания.

Советы и рекомендации

- Для того, чтобы приступить к выполнению задания, необходимо используя образ Mongo, аналогично тому как показано на уроке, установить в Docker образ сервера БД.
- Для установки Docker на Windows 10 и 11 воспользуйтесь инструкциями:
 - Установить Docker на Windows 10 - <https://tretyakov.net/post/ustanovit-docker-na-windows-10-wsl2/>
 - Установить Docker на Windows 11 - <https://youtu.be/6k1CyA5zYgg>

